



CCADET

**CENTRO DE CIENCIAS APLICADAS Y DESARROLLO TECNOLÓGICO
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

Informe técnico

Proceso de evaluación de usabilidad con usuarios

Versión 2012

Gustavo de la Cruz Martínez

Ana Libia Eslava Cervantes

Ricardo Castañeda Martínez

Jesús Ramírez Ortega

Octubre 2012

Tipo de proyecto: Desarrollo

RESUMEN

La interfaz del usuario es el medio a través del cual un usuario humano y una computadora pueden comunicarse e interactuar. Las evaluaciones de usabilidad de un sistema interactivo tendrán como objetivo calificar el impacto de la interfaz sobre la actividad del usuario. La evaluación de usabilidad de un sistema interactivo nos permite identificar con anticipación alguna de los problemas que el usuario puede encontrar al realizar sus actividades con un software interactivo.

Las evaluaciones de usabilidad del grupo ESIE (Espacios y Sistemas Interactivos para la Educación) se centran en analizar la utilidad del sistema, identificar problemas de usabilidad y sugerir soluciones para subsanar esas deficiencias. Estas evaluaciones se realizan en el laboratorio de usabilidad ubicado en las instalaciones del centro de Ciencias Aplicadas y Desarrollo Tecnológico (CCADET) de la UNAM.

Las observaciones que se generan como consecuencia de estas evaluaciones buscan mejorar una propuesta interactiva y con ello apoyar de mejor manera al usuario con la realización de sus actividades.

Índice

Introducción.....	4
1. La interfaz de usuario.....	5
2. Evaluación de la Usabilidad.....	6
2.1 Usabilidad.....	6
2.2 Evaluación de la usabilidad.....	8
2.3 Tipos de evaluación.....	9
2.4 Pruebas de usuarios.....	9
2.5 Evaluación heurística.....	10
2.6 Evaluación automática.....	10
3 proceso de evaluación de usabilidad con usuarios.....	11
3.1 Objetivos.....	11
3.2 Preparación.....	11
3.3 Realización.....	11
3.4 Análisis.....	12
Conclusiones.....	13
Referencias Bibliográficas.....	14

Introducción

En la evaluación de una interfaz de un sistema se determina la capacidad de la interfaz para implementar correctamente todas las tareas del usuario, la facilidad de uso y aprendizaje de la interfaz, y, finalmente, la aceptación por el usuario (Treu, 1998).

La evaluación abarca diversas herramientas que van desde la "prueba de manejo" informal donde el usuario proporciona retroalimentación informal, hasta estudios basados en métodos estadísticos para la evaluación de cuestionarios aplicados a usuarios finales.

El proceso de evaluación muchas veces indica cambios en el diseño de la interfaz, por tanto luego de una evaluación se hacen modificaciones en la interfaz que deberán ser evaluadas nuevamente. Este ciclo puede continuar hasta que ya no sean necesarias más modificaciones a la interfaz.

La evaluación de la interfaz se puede ir haciendo aplicando criterios de evaluación de usabilidad, en las primeras fases del diseño para reducir los tiempos de desarrollo de la interfaz, pero siempre serán necesarias las evaluaciones con los usuarios.

En este documento se describe el proceso de evaluación de usabilidad con usuarios que sigue el grupo de Espacios y Sistemas Interactivos para la Educación ESIE.

1. La interfaz de usuario

La interfaz de usuario es el medio a través del cual un usuario humano y una computadora pueden comunicarse e interactuar. Por tal razón, es necesario proveer al usuario de interfaces que faciliten esta comunicación e interacción; adicionalmente, se requieren interfaces que permitan entender y apoyar a la realización de la tarea de las personas, a partir de las funciones disponibles en el sistema.

Para el diseño de la interfaz de usuario se consideran las necesidades y cualidades del tipo de persona que interactuará con la interfaz del sistema de cómputo, conocido como usuario final. Debido a que se considera el punto de vista del usuario promedio y se define un único comportamiento para todos los usuarios, en la mayoría de las interfaces existe una falta de personalización y manejo de diferencias entre usuarios.

El proceso general de diseño de una interfaz de usuario incluye la definición de las funciones del sistema orientadas a la tarea del usuario, el uso de lineamientos generales para la conceptualización de la interfaz y la evaluación con los usuarios.

Para la definición de la tarea se producen diferentes modelos que ayudan al entendimiento del sistema: modelo del usuario, modelo del diseño, modelo mental y modelo de la implementación (Pressman, 2009). Con el modelo del usuario se busca entender las características (edad, sexo, habilidades físicas, educación, antecedentes culturales o étnicos, motivaciones, objetivos, personalidad) o perfil del usuario que realizará la tarea o usuario final.

En el modelo del diseño se especifican datos, arquitectura y funcionamiento específico (procedimientos) para resolver la tarea. El modelo mental es la percepción del sistema por parte del usuario final. El modelo de la implementación debe combinar la información del usuario y de la tarea (reglas y definiciones) buscando que este modelo coincida con el modelo mental.

Generalmente se considera que el proceso de desarrollo de una interfaz de usuario es iterativo y se representa con un modelo espiral. Este proceso tradicionalmente involucra cuatro actividades (Mandel, 1997):

- 1 Análisis y modelado de usuarios, tareas y el ambiente.
- 2 Diseño de la interfaz.
- 3 Construcción de la interfaz.
- 4 Evaluación de la interfaz.

En cada iteración cada una de las actividades se van refinando.

En la actividad de análisis y modelado se define el modelo del usuario y las características físicas del ambiente de trabajo, con esto se crea un modelo del análisis de la interfaz para continuar con el diseño de la misma. En el diseño se definirán los objetos y las acciones necesarias para que el usuario realice sus tareas. En la mayoría de las ocasiones la construcción de la interfaz parte de la creación de prototipos que permitan la evaluación de aspectos específicos de la interfaz de usuario. En la evaluación se determina la capacidad de la interfaz para implementar correctamente todas las tareas del usuario, la facilidad de uso y aprendizaje de la interfaz, y, finalmente, la

aceptación por el usuario

2. Evaluación de la Usabilidad

2.1. Usabilidad

El estándar ISO 9241 (Ergonomics requirements for office work with visual display terminals – VDTs) (ISO, 1993), describe los requisitos ergonómicos para trabajo de oficina con pantallas de despliegue visual y explica algunos de los principios básicos subyacentes.

El estándar ISO 9241-11 (Guidance on Usability) (ISO, 1998) define como especificar y medir la usabilidad de productos y aquellos factores que tienen un efecto en la usabilidad. La usabilidad de acuerdo a este estándar es: “La extensión para la que un producto puede ser usado por usuarios específicos, para lograr metas específicas con efectividad, eficacia y satisfacción en un contexto de uso específico”.

Para especificar o medir la usabilidad es necesario identificar las metas y descomponer la Efectividad, eficiencia y satisfacción, así como los componentes del contexto de uso en Subcomponentes con atributos medibles y verificables:

- Eficacia: definido en términos de la exactitud y completitud con que usuarios específicos pueden lograr metas específicas en ambientes particulares.
- Eficiencia: referido a los recursos gastados en relación con la precisión y completitud de la meta lograda, es decir recursos de tiempo, financieros y humanos.
- Satisfacción: que evalúa el confort o comodidad y la aceptabilidad del trabajo del sistema para sus usuarios y otras personas afectadas por su uso.

ISO 9241 define la usabilidad en términos de la calidad del trabajo de un sistema en uso, la cual depende de todos los factores que pueden influenciar el uso de un producto en el mundo real: factores organizacionales (prácticas de trabajo, ubicación o apariencia de un producto), diferencias individuales entre usuarios (factores culturales y preferencias), experiencia, etc.

De acuerdo al estándar ISO/IEC 9126 (Software Product Evaluation - Quality Characteristics and Guidelines for the User) (ISO, 1991), usabilidad es un atributo de la calidad del software. El término es utilizado para referirse a la capacidad de un producto para ser usado fácilmente. Esto corresponde a la definición de usabilidad como parte de la calidad del software, siendo la calidad del software definida por el estándar como: “Un conjunto de atributos de software que se sostienen en el esfuerzo necesitado para el uso y en la valoración individual de tal uso por un conjunto de usuarios declarados o implicados”. Esto está relacionado con la capacidad del producto software para ser entendido, aprendido, usado y atractivo para el usuario, cuando es utilizado bajo condiciones específicas.

En la parte 1 del estándar ISO 9126, la usabilidad es analizada en términos de su comprensibilidad, aprendizaje, tolerabilidad, interactividad y complacencia, tal como lo describe Alva (2005):

- Comprensibilidad, define la capacidad del producto software para permitir al usuario entender si el software es adecuado, y como puede ser usado para tareas y condiciones de uso particulares.

- Aprendizaje, referido a la capacidad del producto software para permitir a los usuarios aprender a usar sus aplicaciones.

- Operatividad, es la capacidad del producto software para permitir al usuario operarlo y controlarlo. Aspectos de conformidad, mutabilidad, adaptabilidad e instalación pueden afectar a la operatividad. También este atributo corresponde a la tolerancia de error, y conformidad con las expectativas del usuario. En un sistema, sobre el que opera un usuario, la combinación de funcionalidad, confiabilidad, usabilidad y eficiencia pueden ser medidas externamente por la calidad de uso.

- Atractivo, es la capacidad del producto software para ser atractivo al usuario. Está referido a los atributos del software pensados para hacer el software más atractivo al usuario, tal como el uso de color y la naturaleza del diseño gráfico.

- Conformidad a estándares y pautas, referido a la capacidad del producto software para adherirse a estándares, convenciones, guías de estilo o regulaciones relacionadas con la usabilidad.

Desde la visión de Nielsen (1993), la usabilidad se define en términos de cinco atributos: aprendizaje, eficiencia, memorización, prevención de error y satisfacción subjetiva.

- Aprendizaje, significa que nuevos usuarios deberían aprender fácilmente a usar el sistema.

- Eficiencia, el sistema debería ser eficiente para uso cuando el usuario ha aprendido a usarlo.

- Memorización, el sistema deberá ser fácil de recordar incluso después de algún periodo sin uso.

- Prevención de error, el sistema deberá tener un bajo porcentaje de error y el usuario deberá fácilmente recuperarse de posibles errores.

- Satisfacción, significa que el sistema debe ser agradable de usar.

2.2. Evaluación de la usabilidad

La evaluación de la usabilidad es un proceso para producir una medida de la facilidad de uso. En la evaluación, hay un objeto que está siendo evaluado y un proceso a través del cual uno o más atributos son juzgados o se les da un valor (Karat, 1997). La evaluación de usabilidad para algunos autores como Mathew (1999), es un estudio empírico con usuarios reales del sistema propuesto, con el propósito de proporcionar retroalimentación en el desarrollo de software durante el ciclo de vida de desarrollo iterativo. El campo de desarrollo de software ha reflejado un creciente interés en pruebas de usabilidad, que son generalmente ejecutadas en los estados de desarrollo de nuevos hardware y software. El concepto de evaluación de usabilidad es para permitir la validación de todos los requisitos, para hacerlo tan útil como sea posible y así aumentar la calidad del producto y la satisfacción del cliente del producto potencial.

La evaluación de la usabilidad, es una de las tareas más importantes que debe emprenderse cuando se desarrolla una interfaz de usuario. Las interfaces pobres pueden, en el ambiente comercial, ahuyentar a clientes potenciales o en el ambiente educativo llevar al fracaso a un aprendiz. En el mundo competitivo de ingeniería de software, una interfaz pobre puede empujar a los usuarios a las manos de la competencia.

Algunos métodos de evaluación pueden requerir un completo laboratorio de usabilidad y otros

Pueden lograrse con poco más que una interacción semi-formal entre el grupo de desarrollo y los usuarios. Incluso con una inversión relativamente pequeña en métodos de usabilidad puede

Obtenerse una mejora significativa de la usabilidad de un sistema de software (Nilson y Loranger, 2006). Analizando lo anterior podemos establecer que el propósito de la evaluación de usabilidad es entre otros:

- 1 Proporcionar retroalimentación para mejorar el diseño
- 2 Valorar qué objetivos de usuarios y organizaciones están siendo logrados
- 3 Monitorizar el uso de productos o sistemas a largo plazo.

Ya en los años 80, se había reconocido que las pruebas de usabilidad fallaban para encontrar los prerequisites, como sugiere Whiteside et.al. (1988), debido principalmente a dos factores:

- Algunos estudios de laboratorios se realizaron en condiciones tan distantes del uso real del sistema que la relación de los datos obtenidos respecto a la vida real obtenidos son irrelevantes y distorsionadas.
- El uso de usuarios no representativos (vistos como sujetos de experimento),

Esto condujo a que en los años noventa ya se hubiera determinado que los dos prerequisites principales para tener resultados válidos y útiles en cualquier evaluación de usabilidad eran (Whiteside et al. 2000):

- Los datos, los cuales debían deducirse de circunstancias que tuvieran validez aceptable
- La aplicación de la técnica de análisis de datos más apropiada para los datos obtenidos.

Por otro lado, para efectos de la evaluación de la usabilidad no solo es de interés el escenario físico y de organización de la evaluación, sino también las características de los usuarios y de las tareas. Ya que una evaluación basada en el usuario puede estudiar sólo un subconjunto de todas las

Posibles tareas que un sistema puede soportar, la evaluación debe estar basada en el estudio de las tareas más representativas, escogidas por su frecuencia o criticidad. Las características del usuario son también importantes en la determinación de la usabilidad, de forma que es fundamental que ésta pueda ser evaluada por un grupo de usuarios representativos de usuarios y no por los propios desarrolladores que poco pueden aportar del uso real.

2.3. Tipos de evaluación

Los tipos de evaluación no se encuentran únicamente referidos a la usabilidad, podrán también aplicarse a la accesibilidad. Además, difieren de unos autores a otros en denominación, número y clasificación. En este caso sólo se describen los más destacados. El fin último será confirmar que las ideas de los expertos están de acuerdo con las necesidades de los usuarios y lo que les gusta.

Aunque los tipos que hemos considerado tienen que ver más con la forma de realizar la evaluación, también habrá que tener en cuenta el momento en que se realice. Con arreglo a este criterio distinguimos entre evaluación formativa (sirve para comprobar que las necesidades de los usuarios se van recogiendo), o de evaluación final (la que se realiza para asegurar el éxito del producto ya acabado).

2.4 Pruebas con usuarios

Este método de evaluación de la usabilidad se basa en la observación y análisis de cómo un grupo de usuarios reales utilizan un sistema interactivo. Se irán anotando todos los problemas de uso reales con los que se encuentren para poder solucionarlos posteriormente.

El experimento estará controlado por el evaluador en un laboratorio de usabilidad dotado de ordenadores con conexión a la red. Consistirá en la observación y registro del comportamiento de los usuarios en tareas previamente acordadas. Por ello es recomendable algún método de grabación de los pasos dados en la realización de las tareas, previa autorización de los usuarios, así como un cuestionario en papel para otro tipo de apreciaciones.

Entre las variables a medir se encuentran el tiempo de realización de la tarea, el tiempo utilizado en recordar la estructura de la página, la satisfacción con la página web o el número de errores.

Los participantes en la evaluación, al menos cinco, no deben contar con más información de la necesaria de forma que actúen con libertad en su interacción con la web. Trabajarán por separado pero es recomendable que durante la prueba consulten todas sus dudas para que el evaluador tenga más datos sobre la usabilidad.

Una vez terminadas las tareas, los usuarios se sentarán con los evaluadores para intercambiar opiniones y poner en común sus dificultades y los puntos que destacan, tanto negativos como positivos.

2.5 Evaluación heurística

La evaluación heurística o por criterios se encuadra dentro de los métodos de inspección de la usabilidad de un sitio web. Los lleva a cabo el experto en usabilidad, y se basan en el recorrido y análisis del sitio. A partir de aquí localizará errores y problemas de diseño. El resultado será un informe que dé cuenta de las mejoras necesarias para que la web sea totalmente usable.

Debido a que está comprobado de diversas personas, aún en base a los mismos principios heurísticos, encontrarán diversos problemas, es recomendable usar un grupo de evaluadores. Dichos evaluadores trabajarán de forma individual para, al final, comunicar sus percepciones mediante informes por escrito en el que enumerarán los problemas y los explicarán de acuerdo a los principios de usabilidad.

Ya que lo que se pretende es evitar repetir los errores en el rediseño del sitio, es conveniente jerarquizar la gravedad de estos errores para ir actuando según lo apremiante que sea. En este sentido habrá que tener en cuenta tres factores:

- La frecuencia con que ocurre el problema.
- El impacto que tiene el problema en los usuarios.
- La persistencia del problema en todo el sitio web.

2.6 Evaluación automática

Este último sistema de evaluación consiste en el uso de software que detecta problemas elementales, como tamaños absolutos de fuentes y de tablas, formato de los textos, tamaño de las páginas, tiempos de descarga o enlaces rotos.

Sin embargo, aunque su ventaja es la rapidez y son considerados un buen punto de partida que ahorran trabajo posterior, este tipo de software no es capaz de detectar cuestiones globales de usabilidad (realmente las más importantes). Éstas sólo pueden ser detectadas por los evaluadores.

Por tanto observamos que el software es útil para todos aquellos aspectos cuantificables, que se pueden considerar de una forma objetiva. Sin embargo, aspectos relacionados con la comodidad, la estética, la legibilidad, la capacidad de actuar fácilmente con el sistema, etc., todos ellos más subjetivos y amplios (como decimos, las verdaderas cuestiones de usabilidad), son difícilmente medibles si no es por medio de personas.

3. Proceso de evaluación de usabilidad con usuarios

3.1 Objetivo

La evaluación de usabilidad de un sistema interactivo tendrá como objetivo calificar el impacto de la interfaz sobre la actividad del usuario.

En general todas las evaluaciones de usabilidad se centran en analizar la utilidad del sistema, proporcionar medios para solucionar los errores y sugerir soluciones para subsanar esas deficiencias.

Adicionalmente se pueden considerar otro tipo de objetivos más específicos que dependerán de las metas de aplicación, los usuarios y expectativas creadas.

3.2. Preparación

En las evaluaciones de usabilidad del ESIE se analizan sólo una parte de la funcionalidad de un sistema interactivo, la cual es definida junto con el equipo de desarrolladores del software en cuestión. En conjunto con este equipo, se analizan y determinan las secciones a evaluar y se definen un conjunto de tareas a observar durante las pruebas con usuarios.

Dadas las características de cada desarrollo, se definen las características de los usuarios participantes en la prueba, quedando clasificados generalmente en dos grupos: Un grupo de usuarios tipo, que no conocen ninguna información sobre el desarrollo del sistema y otro grupo, usuarios de control, que conocen algunos detalles sobre el desarrollo.

3.3. Realización

Las pruebas de usabilidad se realizan en el laboratorio de usabilidad ubicado en las instalaciones del Centro de Ciencias Aplicadas y Desarrollo Tecnológico (CCADET) de la UNAM.

Para la realización de la prueba se define un calendario de sesiones, en las que se distribuyen los usuarios participantes. De acuerdo con Nielsen (2000), en la evaluación deben participar al menos 5 usuarios tipo. Durante el desarrollo de la evaluación se graba el audio y video de la interacción de los usuarios con el sistema como apoyo para el análisis de la usabilidad.

Estas grabaciones permiten analizar con detalle el trabajo de cada usuario con el sistema y compararlo respecto al desempeño esperado. Esta comparación permite obtener datos objetivos y fidedignos sobre los problemas de usabilidad que presenta el sistema. También se aplican cuestionarios antes y después de la prueba a los participantes, para conocer tanto el perfil de los usuarios, y con ello asegurar que cada uno cumplía con las características de los usuarios requeridos para el sistema; así como conocer la opinión de los usuarios con respecto a la organización de los elementos en la pantalla, la legibilidad de las letras, la terminología entre otros.

En total, para poder realizar la evaluación del sistema es necesario generar siete instrumentos, que nos permiten analizar la usabilidad del sistema:

- 1 Protocolo de Bienvenida
- 2 Actividades de la prueba
- 3 Fichas de usuario

- 4 Cuestionario de perfil del usuario
- 5 Cuestionario de usabilidad
- 6 Cuestionario de satisfacción del usuario
- 7 Instrumento de evaluación

3.4. Análisis

Una vez terminada la evaluación, se recogen las impresiones de los usuarios por escrito para una comparación con las observaciones de los evaluadores.

Los usuarios responden unos cuestionarios QUIS, el cual está centrado en evaluar su percepción de la usabilidad de la interfaz. Los aspectos que se evalúan con este cuestionario son los siguientes:

- 1 Reacción global al sistema, para ello emplea preguntas basadas en una escala de valoración de nueve puntos de escala likert. En esta sección, no se tratan características específicas de la interfaz, ni la interacción.
- 2 Factores de pantalla, esto hace referencia a las características del nivel léxico de la interfaz como por ejemplo las fuentes, su organización, el control del usuario y la recuperación de errores.
- 3 Terminología y retroalimentación, permite medir la inteligibilidad de los mensajes relacionados con la interacción.
- 4 Factores de aprendizaje, esta sección está diseñada para medir la experiencia de aprendizaje y secuencia lógica de la interacción.
- 5 Capacidades del sistema, se pregunta sobre la experiencia del usuario con respecto a la velocidad, la respuesta, la confiabilidad, el ruido, manejo de errores y la flexibilidad del sistema.

A partir de todos estos datos podemos identificar la forma en que el usuario percibe la interfaz de usuario.

Conclusiones

La evaluación de usabilidad de un sistema interactivo nos permite identificar con anticipación alguna de los problemas que el usuario puede encontrar al intentar trabajar con un software interactivo.

Una retroalimentación visual y acomodo adecuado de los elementos de navegación permitirían al usuario navegar por el sistema con fluidez y sin pérdida de tiempo; comprender dónde se encuentra o que acción es la que debe realizar para continuar con su actividad. Estos son algunos ejemplos de las observaciones que el proceso de evaluación nos permite identificar.

En algunas evaluaciones se han podido hacer algunas observaciones más específicas como aquellas relacionadas con el tiempo de realización de la tarea: “el tiempo necesario para realizar las actividades con el sistema podría mejorar si la organización de las tareas es más consistente con la forma en que los usuarios hacen su tarea de manera convencional”. También se identifica que se ha observado que se pueden aprovechar las habilidades de los usuarios con el manejo de la computadora para facilitarles su trabajo con el sistema.

Todas estas observaciones habrán de servir como medio extraer una serie de conclusiones en términos de deficiencias detectadas por usuarios y evaluadores, así como mejoras propuestas con el fin de mejorar la usabilidad.

Referencias Bibliográficas

ISO 9126: Software product evaluation - Quality characteristics and guidelines for their use. (1991)

ISO 9241: Ergonomic requirements for office work with visual display terminals (parts 1 to 17) (1993).

ISO CD 9241-11: Guidelines for specifying and measuring usability. (1993)

ISO DIS 9241-10: Dialogue principles. (1993)

ISO DIS 9241-14: Menu dialogues. (1993)

ISO/IEC 9126: Software Engineering – Product quality. Part 1: 2001 – Parts 2 to 4: International Organization for Standardization, Geneva.

Karat, J.: User-centered software evaluation methodologies. In M. Helander, T.K.

Landauer, P. Prabhu (Eds.) Handbook of human-computer interaction, 2nd ed. Elsevier Science. (1997)

Mandel, T. 1997 The Elements of User Interface Design. John Wiley & Sons, Inc.

Mayhew, D.: The Usability Engineering Lifecycle. Morgan Kaufmann, San Francisco, California. . (1999)

Nielsen, J. y Loranger, H. 2006 Prioritizing Web Usability. New Riders Publishing.

Nielsen, J.: Why you only need to test with 5 users. Alertbox, (2000) Nielsen (<http://www.alertbox.com/>, 2000)

Nielsen J.: Usability engineering. Academic Press. (1993)

Pressman, R. 2009. Software Engineering: A Practitioner's Approach. 7a ed. McGraw-Hill, Inc., New York, NY, USA.

Treu, S. 1994. User interface evaluation: a structured approach. Plenum Press, New York.

Whittaker, S.; Terveen, L.; Nardi, B.: Let's stop pushing the envelope and start addressing it: A reference task agenda for HCI. Human-Computer Interaction, 15(2/3), 75-106. (2000).

Whiteside, J.; Bennett, J.; Holzblatt, K. 1988. Usability engineering: Our experience and Evolution. In: M Helander (ed.) Handbook of Human-Computer Interaction. Elsevier, 791-817